The Single Allocation Problem
in the Interacting Three-Hub Network

Jinhyeon Sohn
Department of Business Administration, Sun Moon University
Asan-si, Chung-Nam 336-840, Korea

Sungsoo Park*
Department of Industrial Engineering, Korea Advanced Institute of Science and Technology,
Taejon 305-701, Korea


Abstract
We consider the single allocation problem in the interacting three-hub network with fixed hub locations. In the single allocation hub network, the hubs are fully interconnected and each nonhub node has to be connected to exactly one of the hubs. The flows between each pair of nodes are sent using the hubs as intermediate switching points. The problem is to find an optimal allocation of nonhub nodes to the hubs which minimizes the total flow cost. We show that the single allocation problem is NP-hard as soon as the number of hubs is three, although the problem in a two-hub system has polynomial time algorithms. This paper provides a mixed integer formulation of the problem and considers the polyhedral properties of it. The formulation can also be used for the single allocation problem with fixed costs for opening links, the 3-terminal cut problem and the 3-processor distribution problem. Computational experiences are reported for data given in the literature and randomly generated problems.

* Corresponding author. E-mail: sspark@cais.kaist.ac.kr
1. INTRODUCTION

Since O'Kelly [14] considered the location problem of interacting hub facilities, a number of studies have been done on the interacting $p$-hub location problem. One type of $p$-hub location problem is the single allocation $p$-hub location problem which can be defined as follows: The locations of nodes (cities), the amounts of flows and the unit costs of direct flows between each pair of nodes are given. Hubs are central facilities designed to act as switching points for internodal flows and are fully interconnected. The flows between any pair of nodes are sent via hubs and there are economies of scale for inter-hub flows. Then, the problem is to find the locations of hubs among the given nodes and to connect each of the remaining nodes to exactly one of the hubs so that the total transportation cost is minimized. However, in real situations, the hub locations are usually fixed for some time interval because of long term lease contracts, equipment at hubs, cost of moving, etc. In this situation, the decision of optimally assigning the nonhub nodes to the hubs is important for efficient operation of the network. In this paper we consider the single allocation problem in the three-hub system with fixed hub locations (Fig. 1) which is a subproblem of the single allocation $p$-hub location problem. A study on the three-hub system for the planar case can be found in Chapter 9 of Fotheringham and O’Kelly [9].

$$>	ext{Insert Fig. 1 here} <<$$

For the $p$-hub location problem with single allocation, O'Kelly [15] provided a quadratic integer formulation and two heuristic algorithms (nearest hub allocation, and nearest or second nearest hub allocation). He presented computational results of the algorithms with 10, 15, 20, 25 nodes and 2, 3, 4 hubs. A number of solution approaches have been proposed for the problem recently: Exchange and clustering heuristics (Klincewicz [11]), tabu search (Klincewicz [12], Skorin-Kapov and Skorin-Kapov [17]), branch and bound procedures (Aykin [1], Ernst and Krishnamoorthy [8]), neural net approaches (Smith et al. [19]). Skorin-Kapov et al. [18]
provided mixed integer formulations for $p$-hub location problems with single and multiple allocation, and presented some computational results for the linear programming relaxations of them. Ernst and Krishnamoorthy [7] provided a new formulation based on a multicommodity concept, and presented a heuristic algorithm based on simulated annealing. O'Kelly et al. [16] provided a computational study on the single and multiple allocation $p$-hub location problems varying several key parameters. Campbell [2] classified hub location problems and presented integer programming formulations for a variety of them. A survey of network hub location problems was presented by Campbell [3]. For the single allocation problem with fixed hubs, Campbell [4] proposed two heuristics (ALLFLO and MAXFLO) based on the solution of the multiple allocation problem. Sohn and Park [20] provided a linear programming formulation for the case with $p=2$. Sohn and Park [21] considered a modified formulation of Skorin-Kapov et al. [18] and the LP relaxation of it for the cases with $p \geq 3$.

Some of the above papers stated that the single allocation problem with fixed hubs is NP-hard without explicit proof since the quadratic assignment problem is NP-hard. However, the structure of the problem is not the same as the quadratic assignment problem although quadratic terms appear in the formulation. Sohn and Park [20] showed that the single allocation problem can be solved in polynomial time when $p=2$. One of the purposes of this paper is to provide the proof that the single allocation problem is NP-hard as soon as the number of fixed hubs is three. We present it by showing that the 3-terminal cut problem known to be NP-hard (refer to Dahlhaus et al. [6]) can be polynomially transformed to the single allocation problem in the interacting three hub network.

In this paper we provide a new mixed integer programming formulation for the single allocation problem in the 3-hub system with fixed hub locations and consider the polyhedral structure of the convex hull of feasible solutions. Sometimes it might be reasonable to consider fixed costs for opening links between nonhub nodes and hubs according to the transportation mode between them. The formulation which is presented here can be also used for the single allocation problem with fixed costs for opening links in the 3-hub network, the 3-terminal cut
problem in [6] and the 3-processor distribution problem in [22]. We tested the strength of the linear programming relaxation using data given in the literature and randomly generated problems.

The remainder of the paper is organized as follows. In Section 2, we present the quadratic integer programming formulation (QIP) of the problem and show the NP-hardness of the problem. In Section 3, we transform (QIP) into a mixed integer programming formulation (MIP), and consider the polyhedral properties of (MIP). Section 4 provides computational experiences for the linear programming relaxation of (MIP). Concluding remarks are given in Section 5.

2. QUADRATIC 0-1 INTEGER FORMULATION AND THE COMPLEXITY OF THE PROBLEM

Suppose that the flows (traffic) between every pair of nodes are known and the locations of three interconnected hubs are given. We assume without loss of generality that the given hub nodes are numbered 1, 2 and 3. We use the following notation.

\( N = \{1, 2, ..., n \} \): the set of nodes.

\( H = \{1, 2, 3\} \): the set of hub nodes, where \( H \subseteq N \).

\( N_0 = N - H \): the set of nonhub nodes.

\( v_{ij} \): amount of flows between nodes \( i \) and \( j \). We assume that, for \( i < j \), \( v_{ij} \) denotes the total bi-directional flows. Therefore, we set \( v_{ij} = 0 \) whenever \( i \geq j \).

\( \mu_i \): amount of flows originating or terminating at node \( i \); thus we have

\[
\mu_i = \sum_{j \in N, j > i} v_{ij} + \sum_{j \in N, j < i} v_{ji}.
\]

\( x_{ik} = 1 \) if node \( i \) is linked to hub \( k \) and 0 otherwise.

\( c_{ik} \): transportation cost of one unit of flow between nonhub node \( i \) and hub \( k \).

\( c_{h} \): transportation cost of one unit of flow between hub 2 and hub 3.
$c_h^2$: transportation cost of one unit of flow between hub 1 and hub 3.

$c_h^3$: transportation cost of one unit of flow between hub 1 and hub 2.

The economies of scale between hubs are reflected in $c_h^i$. In the fixed hub system, the economies of scale between hubs need not be identical. We assume without loss of generality that all the inter-hub costs are positive and satisfy the triangular inequality, i.e.,

$$c_h^1 < c_h^2 + c_h^3, \quad c_h^2 < c_h^1 + c_h^3, \quad c_h^3 < c_h^1 + c_h^2.$$

Then, the single allocation problem with three fixed hubs can be formulated as the following quadratic 0-1 integer program.

- **Quadratic Integer Program**

\[
\begin{align*}
\text{(QIP)} : \text{Min} & \quad \sum_{i=1}^{n} \sum_{k=1}^{3} \mu_i c_{ik} x_{ik} + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} v_{ij} \left\{ c_h^1 (x_{i1} x_{j3} + x_{i3} x_{j2}) + c_h^2 (x_{i1} x_{j3} + x_{i3} x_{j1}) + c_h^3 (x_{i1} x_{j2} + x_{i2} x_{j1}) \right\} \\
\text{s.t.} & \quad \sum_{k=1}^{3} x_{ik} = 1 \text{ for all } i \in N \\
& \quad x_{11} = x_{22} = x_{33} = 1 \\
& \quad x_{ik} \in \{0, 1\} \quad i \in N, k \in H
\end{align*}
\]  

The first term in (1) represents the sum of flow costs between nonhub node $i$ and hub $k$ for all $i \in N_0, k \in H$. The second term represents the inter-hub flow costs. If node $i$ and node $j$ are assigned to different hubs, the flow cost $c_h^i v_{ij}$ is incurred. Constraints (2) and (4) ensure that each node is assigned to one and only one hub. Constraints (3) imply that nodes 1, 2, and 3 are hub nodes. The above formulation is the same as the formulation in [20] except that the number of hubs is three in the present formulation.

Now, although the single allocation problem with two fixed hubs has a polynomial time algorithm as shown in [20], we show that the single allocation problem with three or more fixed
hubs is NP-hard by showing that the \( k \)-terminal cut problem which is known to be NP-hard for any fixed \( k \geq 3 \) (refer to [6]) can be polynomially transformed to the single allocation problem with \( k \) fixed hubs. We got this idea from the transformation of the single allocation problem with two fixed hubs into the minimum cut problem in [20]. The \( k \)-terminal cut problem is defined as follows. Given an undirected graph \( G = (V, E) \), a set \( S = \{s_1, s_2, \ldots, s_k\} \) of \( k \) specified terminals, \( S \subseteq V \), and a positive weight \( l(e) \) for each edge \( e \in E \), find a minimum weight set of edges \( E' \subseteq E \) such that the removal of \( E' \) from \( E \) disconnects each terminal from all the others in \( S \).

**Theorem 1.** The single allocation problem of fixed \( p \)-hub system is NP-hard for every fixed \( p \geq 3 \).

**Proof.** We show that the 3-terminal cut problem can be polynomially transformed to the single allocation problem with three fixed hubs, and the technique can be readily generalized to cover the cases with general fixed \( p \). For the given undirected graph \( G = (N, E) \) of the 3-terminal cut problem, assume that the three terminal nodes are numbered as 1, 2, 3 and other nodes are numbered as 4, 5, \ldots, \( |N| \). Let \( l(i, j) \) be the weight of the edge connecting nodes \( i \) and \( j \), and if there is no edge between nodes \( i \) and \( j \) let \( l(i, j) = 0 \). To transform the 3-terminal cut problem on \( G \) into the allocation problem of the 3-hub network, let us define the terminal nodes 1, 2, and 3 in \( G \) to be the hubs in the problem that will be constructed, and let \( c_h^1 = c_h^2 = c_h^3 = c_h \) for arbitrary positive number \( c_h \). Let \( v_{ij} = l(i, j) / c_h \) for \( i, j \in N, \ i < j, \ v_{ii} = 0 \) for \( i, j \in N, \ i > j \), and \( c_{ik} = 0 \) for \( i \in N_0, \ k \in H \). Given a feasible solution to the transformed 3-hub system, let \( H_k \) be the set of hub \( k \) and nonhub nodes allocated to hub \( k \). Then, the objective value of the solution to the transformed 3-hub system is
\[ \sum_{i \in H_1, j \in H_2} (v_{ij} + v_{ji}) c_h + \sum_{i \in H_1, j \in H_3} (v_{ij} + v_{ji}) c_h + \sum_{i \in H_2, j \in H_3} (v_{ij} + v_{ji}) c_h = \sum_{i \in H_1, j \in H_2} l(i, j) + \sum_{i \in H_1, j \in H_3} l(i, j) + \sum_{i \in H_2, j \in H_3} l(i, j) \]

We can think that the nodes allocated to the same hub in the solution to the transformed 3-hub system are clustered in a partition of \( G \) and vice versa. Then, an optimal solution to the transformed single allocation problem with three fixed hubs is an optimal solution to the 3-terminal cut problem on \( G \).

The values of \( v_{ij} \) were set to satisfy the equation \( v_{ij} c_h = l(i, j) \) for \( i < j \). The rational behind the transformation can be stated simply in the following way: If node \( i \) and node \( j \) are assigned to different hubs (or clustered to different subsets), the inter-hub flow cost \( v_{ij} c_h \) (or the cut weight \( l(i, j) \)) occurs.

3. MIXED INTEGER PROGRAM

In this section we transform (QIP) into a 0-1 mixed integer program, and consider polyhedral properties and the linear programming relaxation of the formulation. A similar approach was used to transform a quadratic 0-1 integer program into a 0-1 integer program in Helme and Magnanti [10].

3.1. TRANSFORMATION OF (QIP) INTO A MIXED INTEGER PROGRAM

Before transforming (QIP) into a mixed integer program (MIP), we delete the fixed variables \( x_{1k}, x_{2k}, x_{3k} \) in (2) and (3) to simplify the analysis of the formulations that will be given. However, that makes the objective function more complicated. Next, we define zero-one variables \( w_{ij}^k \) for all \( k = 1, 2, 3 \) and \( i, j \in N_0 \) with \( i < j \) as follows:

\[ w_{ij}^1 = x_{i2} x_{j3} + x_{i3} x_{j2}, \quad w_{ij}^2 = x_{i1} x_{j3} + x_{i3} x_{j1}, \quad w_{ij}^3 = x_{i1} x_{j2} + x_{i2} x_{j1}. \quad (5) \]
Then, we replace the quadratic term in the objective function (1) with the variables $w^k_{ij}$. For all $i, j \in N_0$ with $i < j$, $w^k_{ij}$ can be interpreted as

\[ w^1_{ij} = 1 \text{ if one of the nodes } i, j \text{ is assigned to hub 2 and the other is assigned to hub 3.} \]

\[ w^2_{ij} = 1 \text{ if one of the nodes } i, j \text{ is assigned to hub 1 and the other is assigned to hub 3.} \]

\[ w^3_{ij} = 1 \text{ if one of the nodes } i, j \text{ is assigned to hub 1 and the other is assigned to hub 2.} \]

\[ w^1_{ij} = w^2_{ij} = w^3_{ij} = 0 \text{ if node } i \text{ and node } j \text{ are assigned to the same hub.} \]

Without the integrality of every $w^k_{ij}$, we can obtain the next 0-1 mixed integer program (MIP) which is equivalent to (QIP).

(MIP) Minimize

\[
(v_{12}c_h^3 + v_{13}c_h^3 + v_{23}c_h^3) + \sum_{i=4}^{n} \sum_{k=1}^{3} \mu_i c_{ik} x_{ik} + \sum_{i=4}^{n} \sum_{j \in N_0, j < i} v_{ij}(c_h^1 w^1_{ij} + c_h^2 w^2_{ij} + c_h^3 w^3_{ij})
\]

\[ + \sum_{i=4}^{n} \left\{ v_{i1}(c_h^2 x_{i1} + c_h^3 x_{i2}) + v_{i2}(c_h^1 x_{i3} + c_h^1 x_{i1}) + v_{i3}(c_h^1 x_{i2} + c_h^3 x_{i1}) \right\} \]

s.t. \[
\sum_{k=1}^{3} x_{ik} = 1, \quad \text{ for all } i \in N_0
\]

\[
\begin{cases}
  x_{i1} - x_{j1} - w^2_{ij} - w^3_{ij} \leq 0, \\
  x_{i2} - x_{j2} - w^1_{ij} - w^3_{ij} \leq 0, \\
  x_{i3} - x_{j3} - w^1_{ij} - w^2_{ij} \leq 0,
\end{cases}
\]

for all $i, j \in N_0$ with $i < j$

\[ w^k_{ij} \geq 0, \quad \text{ for all } k \in H \text{ and } i, j \in N_0 \text{ with } i < j \]

\[ x_{ik} \in \{0, 1\}, \quad \text{ for all } i \in N_0, k \in H \]

The first term in (6) represents the inter-hub flow costs of the pure traffic demands only between hubs, which is a constant. The second term represents the sum of flow costs between nonhub node $i$ and hub $k$ for all $i \in N_0, k \in H$. The third term represents the inter-hub flow
costs of the traffic demands between nonhub nodes, and the final term represents the inter-hub flow costs of the traffic demands between nonhub nodes and hubs. If there exists fixed cost $f_{ik}$ for opening a link between nonhub node $i$ and hub $k$, it suffices to change the second term of (6) to $\sum_{i=1}^{n} \sum_{k=1}^{3} (f_{ik} + \mu_{i} c_{ik})x_{ik}$. This will not change the validity of the formulation.

Now, we show the equivalence of (QIP) and (MIP). If $v_{lm}$ is zero for some $l, m \in N_0$ with $l < m$, deleting constraints in (8) for every $w_{lm}^{k}$ does not affect the optimal value. From now on, consider the variables $w_{ij}$ only in the case where $v_{ij}$ is positive.

**Lemma 1.** If triangular inequalities hold among positive values $c_{h}^{1}, c_{h}^{2}$ and $c_{h}^{3}$, for the given nonnegative numbers $t_{1}, t_{2}$, and $t_{3}$, the optimal solution of Min $c_{h}^{1}w_{1} + c_{h}^{2}w_{2} + c_{h}^{3}w_{3}$ with the constraints $w_{2} + w_{3} \geq t_{1}$, $w_{1} + w_{3} \geq t_{2}$, $w_{1} + w_{2} \geq t_{3}$ must satisfy the constraints at equality, so the optimal solution is $w_{1} = (t_{2} + t_{3} - t_{1})/2$, $w_{2} = (t_{1} + t_{3} - t_{2})/2$, $w_{3} = (t_{1} + t_{2} - t_{3})/2$.

**Proof.** Assume that there is an optimal solution such that $w_{2} + w_{3} > t_{1}$, and let $w_{2} + w_{3} - t_{1} = \delta$. If we decrease the values of $w_{2}$ and $w_{3}$ by $\delta/2$ respectively and increase the value of $w_{1}$ by $\delta/2$, that will not violate the constraints. Then, the objective will decrease by the amount of $(c_{h}^{2} + c_{h}^{3} - c_{h}^{1})\delta/2$ which is positive. This is a contradiction to the optimality of the given solution. Similar reasoning also holds for the other cases. \(\square\)

**Proposition 1.** (QIP) and (MIP) are equivalent.

**Proof.** Considering the incidence vector of $(x_{i1}, x_{i2}, x_{i3}, x_{j1}, x_{j2}, x_{j3})$ for $i, j \in N_0$ with $i < j$, possible instances satisfying constraints (7) and (10) are as follows:
i) when \(i\) and \(j\) are assigned to the same hub; \((1, 0, 0, 1, 0, 0)\) or \((0, 1, 0, 0, 1, 0)\) or \((0, 0, 1, 0, 0, 1)\),

so we have the constraints \(w_{ij}^2 + w_{ij}^3 \geq 0,\ w_{ij}^1 + w_{ij}^3 \geq 0\), and \(w_{ij}^1 + w_{ij}^2 \geq 0\) from (8).

ii) when one is assigned to hub 1 and the other is assigned to hub 2; \((1, 0, 0, 0, 1, 0)\) or\n
\((0, 1, 0, 1, 0, 0)\), so we have \(w_{ij}^2 + w_{ij}^3 \geq 1,\ w_{ij}^1 + w_{ij}^3 \geq 1,\ w_{ij}^1 + w_{ij}^2 \geq 0\).

iii) when one is assigned to hub 1 and the other is assigned to hub 3; \((1, 0, 0, 0, 0, 1)\) or

\((0, 0, 1, 1, 0, 0)\), so we have \(w_{ij}^2 + w_{ij}^3 \geq 1,\ w_{ij}^1 + w_{ij}^3 \geq 0,\ w_{ij}^1 + w_{ij}^2 \geq 1\).

iv) when one is assigned to hub 2 and the other is assigned to hub 3; \((0, 1, 0, 0, 0, 1)\) or

\((0, 0, 1, 0, 1, 0)\), so we have \(w_{ij}^2 + w_{ij}^3 \geq 0,\ w_{ij}^1 + w_{ij}^3 \geq 1,\ w_{ij}^1 + w_{ij}^2 \geq 1\).

Since the objective function's coefficients of every \(w_{ij}^k\) (for \(v_{ij} > 0\)) are positive and satisfy the triangular inequalities, we can see that the values of \(w_{ij}^k\) will be as follows at optimality from Lemma 1;

\[
\begin{align*}
\text{i) } & w_{ij}^1 = w_{ij}^2 = w_{ij}^3 = 0, & \text{ii) } & w_{ij}^3 = 1, w_{ij}^1 = w_{ij}^2 = 0, \\
\text{iii) } & w_{ij}^2 = 1, w_{ij}^1 = w_{ij}^3 = 0, & \text{iv) } & w_{ij}^1 = 1, w_{ij}^2 = w_{ij}^3 = 0.
\end{align*}
\]

(MIP) has \(3(n-3)\) \(0/1\) variables, \(3(n-3)(n-4)/2\) continuous variables, and \((n-3)+3(n-3)(n-4)\) constraints. Compared to the formulation with fixed hubs in [21], the number of continuous variables is reduced to a third.

As remarked in [20], the \(p\)-processor distribution problem considered in [22] has the same cost structure as the single allocation problem with \(p\)-fixed hubs except that there is no pure flow demand between processors. In the multiprocessor distribution problem, let \(a_{ik}\) denote the cost of computation of module \(i\) on processor \(k\) and \(c_{ijpq}\) denote the interprocessor communication
cost if module \(i\) is assigned to processor \(p\) and module \(j\) is assigned to processor \(q\). If we delete the first constant term in (6) of (MIP), and use \(a_{ik}\) instead of \(\mu, c_{ik}\) and \(c_{ijpq}\) instead of \(v_{ij}c_{h}^p\) in (MIP), the formulation can be used to solve the 3-processor scheduling problem when triangular inequalities hold between the interprocessor communication costs.

### 3.2. POLYHEDRAL PROPERTIES OF (MIP)

In this subsection we show that every inequality in (8) defines a facet of the convex hull of feasible solutions of (MIP). We may expect that the LP relaxation which uses the facet-defining inequality gives a tight lower bound. For the background material on the polyhedral theory, we refer the reader to Nemhauser and Wolsey [13].

Let \(S_{MIP}\) denote the set of feasible solutions to problem (MIP), and \(CHMIP\) denote the convex hull of \(S_{MIP}\), i.e.,

\[
CHMIP = \text{conv}\left\{(x, w) \in \{0,1\}^{3(n-3)} \times \mathbb{R}^{3(n-3)(n-4)/2} : (x, w) \text{ satisfying (7), (8)}\right\}.
\]

**Proposition 2.** \(\dim (CHMIP) = (3n^2 - 17n + 24)/2\).

**Proof.** Since every feasible solution \((x, w) \in S_{MIP}\) satisfies constraint (7), \(\dim (CHMIP) \leq 3(n-3) + 3(n-3)(n-4)/2 - (n-3) \equiv (3n^2 - 17n + 24)/2\). To prove that the equality holds we show that constraints in (7) are the only constraints that hold at equalities for every feasible solution. Assume that every feasible solution satisfies the constraint

\[
\sum_{i \in N_0} \sum_{k \in H} \alpha_{ik}x_{ik} + \sum_{k=1}^{3} \sum_{i \in N_0} \sum_{j > i} \beta_{ij}^{k}w_{ij}^{k} = \gamma_0. \tag{11}
\]
We will show that the constraint (11) can be expressed as a linear combination of constraints (7), i.e., \( \sum_{i \in N_0} \sum_{k \in H} x_{ik} = \sum_{i \in N_0} \alpha_i. \)

Consider a feasible solution such that every nonhub node is assigned to hub 1 and every \( w_{ij}^k \) is zero. Then, we have that \( \sum_{i \in N_0} \alpha_i = \gamma_0 \) from (11). Similarly, consider a set of feasible solutions such that every nonhub node is assigned to hub 1 and every \( w_{ij}^k \) is zero except that \( w_{pq}^t = 1 \) for some \( p, q \in N_0 \) with \( p < q \), and \( t \in H \). Then, (11) becomes \( \sum_{i \in N_0} \alpha_{i1} + \beta_{pq}^t = \gamma_0 \), so, we see that \( \beta_{pq}^t = 0 \). Since \( p, q, \) and \( t \) are arbitrary, every \( \beta_{ij}^k = 0 \). Now, consider a set of feasible solutions such that a nonhub node \( p \) is assigned to hub 2 and the other nonhub nodes are assigned to hub 1. Then, (11) becomes \( \sum_{i \in N_0, j \neq p} \alpha_{i1} + \alpha_{p2} = \gamma_0 \). So, we have that \( \alpha_{p1} = \alpha_{p2} \) for all \( p \in N_0 \). By similar arguments, we can have \( \alpha_{i1} = \alpha_{i2} = \alpha_{i3} = \alpha_i \) for all \( i \in N_0 \). So we have that \( \sum_{i \in N_0} \sum_{k \in H} x_{ik} = \sum_{i \in N_0} \alpha_i (\equiv \gamma_0). \)

\[ \square \]

**Proposition 3.** Every inequality in (8) of the form \( x_{it} - x_{jt} - \sum_{k \in H - \{t\}} w_{ij}^k \leq 0 \) and \(- x_{it} + x_{jt} - \sum_{k \in H - \{t\}} w_{ij}^k \leq 0 \) defines a facet of \( CHMIP \).

**Proof.** We show the result for the form \( x_{i1} - x_{j1} - w_{ij}^2 - w_{ij}^3 \leq 0 \). Assume that every feasible solution \((x, w) \in S_{MIP}\) with \( x_{p1} - x_{q1} - w_{pq}^2 - w_{pq}^3 = 0 \) for some \( p, q \in N_0 \) with \( p < q \), satisfies

\[
\sum_{i \in N_0} \sum_{k \in H} \alpha_{ik} x_{ik} + \sum_{k \in H} \sum_{i \in N_0} \sum_{j \in N_0, j > i} \beta_{ij}^k w_{ij}^k = \gamma_0.
\]
We will show that (12) can be expressed as a linear combination of constraints in (7) and
\begin{equation}
x_{p1} - x_{q1} - w_{pq}^2 - w_{pq}^3 = 0, \text{ i.e., } \sum_{i \in N_0} \alpha_{i_1} \sum_{k \in H} x_{ik} + \rho(x_{p1} - x_{q1} - w_{pq}^2 - w_{pq}^3) = \sum_{i \in N_0} \alpha_i.
\end{equation}

As in the proof of Proposition 2, we have that \( \sum_{i \in N_0} \alpha_{ik} = \gamma_0 \) for \( k \in H \), \( \alpha_{i_1} = \alpha_{i_2} = \alpha_{i_3} \) for \( i \in N_0 - \{p, q\} \), and \( \beta_k^j = 0 \) for all \( i, j \in N_0 \) with \( i < j \) and \( k \in H \) except \( \beta_{pq}^2 \) and \( \beta_{pq}^3 \).

Consider feasible solutions such that nodes \( p, q \) are connected to the same hub 2 or hub 3 with \( w_{pq}^2 = w_{pq}^3 = 0 \) and the other nodes are connected to hub 1. Then, (12) becomes
\begin{equation}
\sum_{i \in N_0 - \{p, q\}} \alpha_{i_1} + \alpha_{p2} + \alpha_{q2} = \gamma_0 \quad \text{or} \quad \sum_{i \in N_0 - \{p, q\}} \alpha_{i_1} + \alpha_{p3} + \alpha_{q3} = \gamma_0,
\end{equation}
so we have that \( \alpha_{p1} + \alpha_{q1} = \alpha_{p2} + \alpha_{q2} = \alpha_{p3} + \alpha_{q3} \).

Consider feasible solutions such that node \( q \) is connected to hub 2 and the other nonhub nodes are connected to hub 1 with \( w_{pq}^1 = w_{pq}^2 = 1, \ w_{pq}^3 = 0 \) or with \( w_{pq}^3 = 1, \ w_{pq}^2 = 0 \). Then, (12) becomes
\begin{equation}
\sum_{i \in N_0 - \{q\}} \alpha_{i_1} + \alpha_{q2} + \beta_{pq}^2 = \gamma_0 \quad \text{or} \quad \sum_{i \in N_0 - \{q\}} \alpha_{i_1} + \alpha_{q2} + \beta_{pq}^3 = \gamma_0,
\end{equation}
so we have that \( \beta_{pq}^2 = \beta_{pq}^3 \) and \( \alpha_{q1} = \alpha_{q2} = \alpha_{q3} \).

Similarly, consider a feasible solution such that node \( q \) is connected to hub 3 and other nodes are connected to hub 1 with \( w_{pq}^2 = 1, \ w_{pq}^3 = 0 \). Then, (12) becomes
\begin{equation}
\sum_{i \in N_0 - \{q\}} \alpha_{i_1} + \alpha_{q3} + \beta_{pq}^2 = \gamma_0, \quad \text{so we have that } \alpha_{q1} = \alpha_{q3} + \beta_{pq}^2.
\end{equation}
Let \( -\rho = \beta_{pq}^2 (= \beta_{pq}^3) \). Then, we have \( \alpha_{q2} = \alpha_{q3} = \alpha_{q1} + \rho, \ \alpha_{p2} = \alpha_{p3} = \alpha_{p1} - \rho, \) and (12) becomes
\begin{equation}
\sum_{i \in N_0} \alpha_{i_1} \sum_{k \in H} x_{ik} - \rho(x_{p2} + x_{p3}) + \rho(x_{q2} + x_{q3}) - \rho(w_{pq}^2 + w_{pq}^3) = \gamma_0. \quad \text{Let } \alpha_i = \alpha_{i_1} \text{ for all } i \in N_0, \text{ since } x_{p2} + x_{p3} = 1 - x_{p1} \text{ and } x_{q2} + x_{q3} = 1 - x_{q1}, \text{ we can obtain}
\begin{equation}
\sum_{i \in N_0} \alpha_{i} \sum_{k \in H} x_{ik} + \rho(x_{p1} - x_{q1} - w_{pq}^2 - w_{pq}^3) = \sum_{i \in N_0} \alpha_i.
\end{equation}

Similar arguments can be used for the other cases.
4. COMPUTATIONAL EXPERIENCE FOR THE LP RELAXATION OF (MIP)

Relaxing the integrality constraints (10), we obtain the following linear programming relaxation of (MIP).

(LPR) : Minimize (6)

with (7), (8), (9),

\[0 \leq x_{ik} \leq 1 \quad \text{for} \quad i \in N, k \in H.\]

Computational experiments using (LPR) were performed on a HP9000/715 workstation with a dual simplex routine of CPLEX Version 3.0 library. We used two classes of data; the data of O'Kelly [15] and randomly generated problems. The data of [15] (also used in [4, 11, 12, 16, 17, 18, 21]) consists of problems with 10, 15, 20 and 25 nodes. For each problem, all instances of \(n \times C_3\) hub location combinations were considered with 10 different values of \(\alpha\) (the economies of scale between hubs) ranging from 0.1 to 1.0 in increments of 0.1.

Random problems having 50 and 100 nodes were created by randomly generating 50 and 100 points respectively in a 10.00 by 10.00 rectangle. We used the randomize and rand functions of the Turbo-C library to generate the pseudorandom numbers. The flows between each pair of nodes were also randomly generated from 0.00 to 10.00, and unit flow costs between them were treated as proportional to the Euclidean distance. Respectively, 100 and 10 randomly selected hub location combinations were tested with the same values of \(\alpha\) as above.

>> Insert Table 1 here <<

Table 1 shows the number of tested instances, the average CPU time for one instance of each problem, and the number of instances providing noninteger solutions. In Table 1, 50-F is
the problem containing fixed costs $f_{ik}$'s for opening links between nonhub nodes and hubs in the above 50 node data. In the 50 node data, means of $\mu_i$ and $c_{ij}$ are 238.95 and 48.54 respectively, so the average flow cost between arbitrary two nodes is 11599. We set fixed costs for opening links between nonhub nodes and hubs as the values in the neighborhood of half the average flow cost or of the average flow cost. The cases considered for 50-F are as follows; for every nonhub node $i$, $f_{i1}$=3000, $f_{i2}$=5000, $f_{i3}$=7000 with 5 values of $\alpha$ ranging from 0.1 to 0.9 in increments of 0.2 and $f_{i1}$=8000, $f_{i2}$=10000, $f_{i3}$=12000 with 5 values of $\alpha$ ranging from 0.2 to 1.0 in increments of 0.2. For each case, 100 hub location combinations used in the above 50 node problem were tested. For the problem with fixed costs, (LPR) gave integer solutions in all instances. Also, the average CPU time for the problem with fixed costs was a little shorter than that for the problem without fixed costs. We think that the fixed costs for opening links may have the effect of fixing some links to be open so that the problem becomes easier to solve. For the above problems, (LPR) gave integer solutions in almost all instances (42241 out of 42250).

To evaluate the performance of (LPR) for the case of the 3-terminal cut problem, the submodularity counter example of Fig. 11 in Dahlhaus et al. [6] and some randomly generated problems were used. 10-C and 50-C consist of 10 and 50 nodes respectively with some edges having positive weights whose values are randomly generated from 1 to 10. The edges having positive weights are randomly selected with probability from 10% to 50% in increments of 10% among all possible edges between every node pair. As terminal nodes, all possible 3-node combinations (in the case of 10-C) and randomly generated 100 combinations of 3-node (in the case of 50-C) were used. We applied (LPR) to the 3-terminal cut problems by transforming the given 3-terminal cut problems into the single allocation problems with three fixed hubs as stated in the proof of Theorem 1, in which $c_{ik} = 0$. It may readily result in fractional solutions because of the cost structure of the transformed problem. However, in our experiments, (LPR) gave integer solutions for all of the transformed problems. There are alternative
transformations to make the values of $c_{ik}$ be non-zero. One of them is as follows: Let $v_{ij} = l(i,j)/c_{h}$ for $i, j \in N_0, i < j$ and for $i, j \in H, i < j$, $v_{ki} = l(k,i)/2c_{h}$ for $k \in H, i < j$, and every other $v_{ij}$ be zero. Let $\mu_i = \sum_{j \in N, j \neq i} v_{ij} + \sum_{j \in N, j < i} v_{ji}$ for $i \in N$, and $c_{i1} = (l(i,2) + l(i,3))/2\mu_i$, $c_{i2} = (l(i,1) + l(i,3))/2\mu_i$, $c_{i3} = (l(i,1) + l(i,2))/2\mu_i$ for $i \in N_0$.

When (LPR) resulted in noninteger solutions, we used the branch and bound method implemented in CPLEX to obtain integer optimal solutions. For all instances, the integer optimal solutions were obtained by branching off only root node on the branch and bound tree. The gaps between the objective values of initial fractional solutions and those of integer optimal solutions were within 0.28%. Table 2 summarizes the computational results, which provides the hub locations yielding fractional solutions for the data of [15]. The same as in [18,21], all noninteger solutions in (LPR) occurred in the cases where $\alpha$ is large ($\alpha \geq 0.7$). In Table 2, the third column denotes the number of nodes having fractional value $x_{ik}$'s for some $k$ in the optimal solution of (LPR). In all noninteger solutions, the value of every $x_{ik}$ having noninteger value was 0.5. Sometimes, e.g. node packing problem, it can be shown that all extreme point solutions of the LP relaxation of integer programming have values in $\{0, 1/2, 1\}$. However, the half-integral property does not hold for our problem.

>> Insert Table 2 here <<

Table 3 provides the computation times taken for solving all instances, with varying $\alpha$, for the problems of O’Kelly [15]. We include it here to compare the performance of our model with other algorithms finding optimal hub locations and optimal allocations together. Generally, computation times for the instances with small $\alpha$ were shorter than those for the instances with large $\alpha$. Computation times for (LPR) were a little shorter than those for the model given in [21], which were about a fourth of the times taken for the model given in [18]. Although
comparisons of running times on different computers are difficult, we think that the algorithm of Ernst and Krishnamoorthy [7] is more efficient than ours in obtaining optimal hub locations. The computation time to find optimal hub locations in our model increases rapidly as the number of nodes increases since we should solve all instances of \( nC_3 \) hub location combinations. However, as shown in Table 1 and 2, computation time of our model was small for each instance. So, our model has merits in solving the subproblem to determine optimal allocation of nonhub nodes under the condition of fixed hub locations. Also, our model can solve the problem with fixed costs for opening links. We think that it can be used to determine hub locations for sufficiently large problems if the number of candidate locations to place the hubs is small.

>> Insert Table 3 here <<

5. CONCLUDING REMARKS

In this paper, we have considered the single allocation problem in the interacting three-hub network. The decision of optimally assigning the nodes to the hubs is important for efficient operation of the hub network with fixed hubs. We have shown that the single allocation problem with \( p \) fixed hubs is NP-hard as soon as the number of hubs is three by showing that the 3-terminal cut problem which is known to be NP-hard can be transformed into our problem. We also provided a 0-1 mixed integer formulation for the single allocation problem with three fixed hubs. The formulation can also be used for the single allocation problem with fixed costs for opening links in the 3-hub network, the 3-terminal cut problem and the 3-processor distribution problem.

We tested the formulation with data given in the literature and randomly generated problems. Randomly generated problems include problems with fixed costs for opening links between nonhub nodes and hubs, and the 3-terminal cut problem. In the computational experiences the
proposed model worked well; only 9 instances out of 43351 instances yielded noninteger solutions in the linear programming relaxation of the provided formulation. For the instances giving noninteger solutions, a commercial branch and bound code could be used. The approach can be used to solve sufficiently large problems when the hub locations are fixed. It can also be used to determine the hub locations if the number of candidate locations for hubs is small.

The problem we have considered here restricts the number of fixed hubs at three. Further research is needed on solving the problem involving more fixed hubs than three. In the fixed hub model, the fixed costs for opening hubs can be treated as a constant term. As remarked in [21], if there is no fixed cost for opening hubs and links, the transportation cost of the problem having more hubs is not higher than that of the problem having less hubs. The problem which simultaneously determines the proper number of hubs, the hub locations and the allocation of nonhub nodes under the condition incurring fixed costs for opening hubs and links also needs further research.

Acknowledgements

The authors are grateful to the referees for the comments that helped us to improve the presentation. We also thank Mohan Krishnamoorthy for sending us the copies of recent papers concerning hub location problems.

APPENDIX

We can think of an alternative mixed integer formulation of the allocation problem using the following simple constraints (A-1) instead of (8),

\[
\begin{align*}
    x_{il} + x_{j2} - w_{ij}^2 & \leq 1, \\
    x_{il} + x_{j3} - w_{ij}^2 & \leq 1, \\
    x_{i2} + x_{j1} - w_{ij}^2 & \leq 1, \\
    x_{i3} + x_{j1} - w_{ij}^2 & \leq 1, \\
    x_{i2} + x_{j3} - w_{ij}^2 & \leq 1, \\
    x_{i3} + x_{j2} - w_{ij}^2 & \leq 1,
\end{align*}
\text{for all } i, j \in N_0 \text{ with } i < j. \quad (A-1)
\]
Constraints (A-1) represent that if node $i$ and node $j$ are allocated to different hubs, then the variable $w_{ij}^k$ which denotes the inter-hub flow has to be 1. Also, in this formulation, the integrality on $w_{ij}^k$ is unnecessary. However, when we consider the linear programming relaxation of each formulation, the formulation using (8) is stronger than that using (A-1), as shown in the next.

Let the mixed integer formulation using (A-1) be (MIP-a) and the linear programming relaxation of (MIP-a) be (LPR-a). We can readily see from Lemma 1 that the optimal solution of (LPR) can be obtained at the point which satisfies at least one of the forms

$$x_{it} - x_{jt} - \sum_{k \in H - \{t\}} w_{ij}^k \leq 0 \quad \text{and} \quad -x_{it} + x_{jt} - \sum_{k \in H - \{t\}} w_{ij}^k \leq 0$$

at equality for every $i, j \in N_0$ with $i < j$, $t \in H$. Let $S_{LPR}$ denote the set of feasible solutions of (LPR) satisfying the above property, and let $z$(LPR) and $z$(LPR-a) denote the optimal objective values of (LPR) and (LPR-a) respectively.

**Proposition A.** $z$(LPR-a) $\leq$ $z$(LPR).

**Proof.** We show that a point in $S_{LPR}$ is a feasible solution of (LPR-a). Given a point in $S_{LPR}$, considering the values of $x_{ik}$ and $x_{jk}$ for arbitrary $i, j \in N_0$ with $i < j$, possible instances are given as follows since $\sum_{k \in H} x_{ik} = \sum_{k \in H} x_{jk} = 1$:

i) $(x_{i1} \geq x_{j1}, x_{i2} \geq x_{j2}, x_{i3} \leq x_{j3})$, 

ii) $(x_{i1} \geq x_{j1}, x_{i2} \leq x_{j2}, x_{i3} \geq x_{j3})$,

iii) $(x_{i1} \geq x_{j1}, x_{i2} \leq x_{j2}, x_{i3} \leq x_{j3})$, 

iv) $(x_{i1} \leq x_{j1}, x_{i2} \geq x_{j2}, x_{i3} \geq x_{j3})$,

v) $(x_{i1} \leq x_{j1}, x_{i2} \geq x_{j2}, x_{i3} \leq x_{j3})$, 

vi) $(x_{i1} \leq x_{j1}, x_{i2} \leq x_{j2}, x_{i3} \geq x_{j3})$. 

19
In the case of i), \( w_{ij}^2 + w_{ij}^3 = x_{i1} - x_{j1}, \) \( w_{ij}^1 + w_{ij}^3 = x_{i2} - x_{j2}, \) \( w_{ij}^1 + w_{ij}^2 = x_{i3} - x_{j3}, \) so we have \( w_{ij}^1 = x_{i2} - x_{j2}, \) \( w_{ij}^2 = x_{i1} - x_{j1}, \) \( w_{ij}^3 = 0. \) Then, these values satisfy (A-1) as follows: It is easy to see that the first term and the first inequalities of the second and the third terms of (A-1) are satisfied. The second inequalities of the second and the third terms of (A-1) are satisfied since \( x_{i3} + x_{j1} - w_{ij}^2 \leq x_{i3} + x_{j1} \leq x_{i3} + x_{j1} \leq 1, \) \( x_{i3} + x_{j2} - w_{ij}^1 \leq x_{i3} + x_{j2} \leq x_{i3} + x_{j2} \leq 1 \) from i). The other cases can be shown similarly. \( \square \)

We can provide a simple example which shows the reverse does not hold. Consider a problem defined as follows: \( H = \{1, 2, 3\}, \) \( N_0 = \{4, 5, 6\}, \) \( c_h^1 = c_h^2 = c_h^3 = 1, \) \( c_{41} = c_{42} = c_{51} = c_{53} = c_{62} = c_{63} = 1, \) \( c_{43} = c_{52} = c_{61} = 1.5, \) \( v_{45} = v_{46} = v_{56} = 1 \) and other \( v_i \)'s are 0. For the instance, (LPR-a) yields an optimal solution such that \( x_{41} = x_{42} = x_{51} = x_{53} = x_{62} = x_{63} = 0.5, \) and all other variables are 0 with \( z(\text{LPR-a}) = 6. \) However, (LPR) yields an optimal solution such that \( x_{41} = x_{51} = x_{61} = 1, \) and all other variables are 0 with \( z(\text{LPR}) = 7. \)
REFERENCES


Fig. 1. A single allocation three-hub network
Table 1. Computational Results for Solving (LPR)

<table>
<thead>
<tr>
<th>Model Type (n)</th>
<th>Number of Tested Instances</th>
<th>Number of Noninteger Solutions</th>
<th>Average CPU Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1,200</td>
<td>0</td>
<td>0.04</td>
</tr>
<tr>
<td>15</td>
<td>4,550</td>
<td>2</td>
<td>0.14</td>
</tr>
<tr>
<td>20</td>
<td>11,400</td>
<td>3</td>
<td>0.44</td>
</tr>
<tr>
<td>25</td>
<td>23,000</td>
<td>2</td>
<td>1.03</td>
</tr>
<tr>
<td>50</td>
<td>1,000</td>
<td>2</td>
<td>23.36</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>0</td>
<td>382.75</td>
</tr>
<tr>
<td>50-F</td>
<td>1,000</td>
<td>0</td>
<td>19.37</td>
</tr>
<tr>
<td>9-Dalhaus</td>
<td>1</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>10-C</td>
<td>600</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>50-C</td>
<td>500</td>
<td>0</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Table 2.
Computational Results for the Instances Yielding Noninteger Solutions in (LPR)

<table>
<thead>
<tr>
<th>Model Type $(n, \alpha, 3 \text{ hubs})$</th>
<th>LP_obj</th>
<th>Number of Noninteger Nodes</th>
<th>IP_obj</th>
<th>% Gap*</th>
<th>CPU Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15, 0.8,4,5,13</td>
<td>1210.66</td>
<td>6</td>
<td>1211.28</td>
<td>0.05</td>
<td>0.5</td>
</tr>
<tr>
<td>15, 1.0,6,11,13</td>
<td>1313.14</td>
<td>6</td>
<td>1316.86</td>
<td>0.28</td>
<td>0.5</td>
</tr>
<tr>
<td>20, 0.7,2,10,15</td>
<td>1215.71</td>
<td>4</td>
<td>1215.89</td>
<td>0.01</td>
<td>0.8</td>
</tr>
<tr>
<td>20, 0.7,2,13,15</td>
<td>1199.08</td>
<td>6</td>
<td>1199.18</td>
<td>0.01</td>
<td>0.8</td>
</tr>
<tr>
<td>20, 1.0,13,15,17</td>
<td>1332.87</td>
<td>6</td>
<td>1333.47</td>
<td>0.05</td>
<td>1.0</td>
</tr>
<tr>
<td>25, 0.8,8,18,24</td>
<td>1358.82</td>
<td>4</td>
<td>1359.00</td>
<td>0.01</td>
<td>1.9</td>
</tr>
<tr>
<td>25, 1.0,3,13,15</td>
<td>1665.74</td>
<td>6</td>
<td>1666.23</td>
<td>0.03</td>
<td>3.1</td>
</tr>
<tr>
<td>50, 0.7</td>
<td>479103.47</td>
<td>10</td>
<td>479103.88</td>
<td>0.00</td>
<td>31.8</td>
</tr>
<tr>
<td>50, 0.8</td>
<td>516212.94</td>
<td>8</td>
<td>516249.75</td>
<td>0.01</td>
<td>37.7</td>
</tr>
</tbody>
</table>

*100×( LP_obj − IP_obj ) / (LP_obj)
Table 3. Computation Times for Solving the data of O’Kelly with varying $\alpha$  (CPU seconds)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>5.0</td>
<td>60.8</td>
<td>470.5</td>
<td>2174.5</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>4.9</td>
<td>61.4</td>
<td>481.2</td>
<td>2231.8</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>4.9</td>
<td>62.0</td>
<td>486.9</td>
<td>2282.3</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>5.2</td>
<td>63.0</td>
<td>495.4</td>
<td>2312.4</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>4.9</td>
<td>63.6</td>
<td>502.6</td>
<td>2384.9</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>4.9</td>
<td>65.1</td>
<td>510.0</td>
<td>2398.7</td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>5.1</td>
<td>64.2</td>
<td>520.3</td>
<td>2418.1</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>5.0</td>
<td>65.2</td>
<td>523.2</td>
<td>2453.3</td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>5.1</td>
<td>65.0</td>
<td>527.7</td>
<td>2473.7</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>4.9</td>
<td>65.1</td>
<td>533.1</td>
<td>2476.8</td>
<td></td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>5.1</td>
<td>63.5</td>
<td>505.1</td>
<td>2360.7</td>
<td></td>
</tr>
</tbody>
</table>